

# Unifying approach for fluctuation theorems from joint probability distributions

Reinaldo García-García,<sup>1</sup> Daniel Domínguez,<sup>1</sup> Vivien Lecomte,<sup>2,3</sup> and Alejandro B. Kolton<sup>1</sup>

<sup>1</sup>*Centro Atómico Bariloche and Instituto Balseiro, 8400 S. C. de Bariloche, Argentina*

<sup>2</sup>*Laboratoire de Probabilités et Modèles Aléatoires (CNRS UMR 7599),*

*Université Pierre et Marie Curie - Paris VI Université Paris-Diderot - Paris VII, 2 Place Jussieu, 75005 Paris, France*

<sup>3</sup>*DPMC-MaNEP, Université de Genève, 24 Quai Ernest-Ansermet, 1211 Geneva 4, Switzerland*

Any decomposition of the total trajectory entropy production for Markovian systems has a joint probability distribution satisfying a generalized detailed fluctuation theorem, when all the contributing terms are odd with respect to time reversal. The expression of the result does not bring into play dual probability distributions, hence easing potential applications. We show that several fluctuation theorems for perturbed non-equilibrium steady states are unified and arise as particular cases of this general result. In particular, we show that the joint probability distribution of the system and reservoir trajectory entropies satisfy a detailed fluctuation theorem valid for all times although each contribution does not do it separately.

PACS numbers: 05.40.-a, 05.70.Ln

The non-equilibrium stochastic thermodynamics of small systems has attracted a lot of attention in the last years. From the experimental side the development of techniques for microscopic manipulation has allowed to study fluctuations in small systems of interest in Physics, Chemistry and Biology [1, 2]. From the theoretical side a group of exact relations known as fluctuation theorems (FT) [3–8] has shed light into the principles governing dissipation and fluctuations in non-equilibrium phenomena, as in driven systems in contact with thermal bath. Formally, the generality of the FTs can be attributed to the way probability distribution functions of different observables behave under time-reversal symmetry-breaking perturbations (see [9, 10] for reviews on FT).

Oono and Paniconi [11] proposed a phenomenological framework for a “non-equilibrium steady-state (NESS) thermodynamics” aimed at describing fluctuating systems subjected to an external protocol. In this approach, the total exchanged heat during a time interval  $\tau$  by a system initially prepared in a NESS is expressed as the sum of two contributions,  $Q^{\text{tot}} = Q^{\text{ex}} + Q^{\text{hk}}$ . The “excess heat”  $Q^{\text{ex}}$  is, in average, associated with the energy exchange during transitions between steady-states while the “housekeeping heat”  $Q^{\text{hk}}$  corresponds, in average, to the energy we need to supply to maintain the system in a NESS. Hatano and Sasa [12] introduced a formal framework for these phenomenological ideas and derived a FT which extends the second law of thermodynamics for transitions between NESS controlled by external parameters  $\sigma(t)$ . The Hatano and Sasa FT is applicable to trajectories  $x(t)$  evolved with a Langevin or more generally a Markovian dynamics starting from an initial condition sampled from a NESS compatible with the initial values  $\sigma(0)$  of the control parameters.

Under identical conditions different FTs were subsequently proposed for NESS, involving the above decomposition of the exchanged heat. We can distinguish between the so-called integral and detailed FTs for sys-

tems initially prepared as described above. The integral fluctuation theorems (IFT) are exact relations for the average over histories of different stochastic trajectory functionals  $W[x]$ , such as  $\langle e^{-W} \rangle = 1$ . Examples are the Jarzynski FT for the total work [7], the Hatano-Sasa FT [12] and the Speck-Seifert FT [13]. The so called detailed FT (DFT) are, on the other hand, exact relations for the probability distributions functions (PDF) of different observables  $W$ , such as  $P(W)/P^{\text{R}}(-W) = e^W$  where  $P^{\text{R}}(W)$  corresponds to the time-reversed protocol  $\sigma^{\text{R}}(t) = \sigma(\tau - t)$ , and a NESS initial condition compatible with  $\sigma(\tau)$ . Examples are given by the Crooks relation [8], or Seifert fluctuation theorem [14]. While observables satisfying a DFT trivially satisfy an IFT, the opposite is not always true. In many recently formulated DFTs, a modified “dual” PDF  $P^{\dagger \text{R}}(W)$  enters into play [15, 16], which corresponds to trajectories generated by the adjoint dynamics (defined below), in general different from the original dynamics of the system. The presence of dual distributions is a strong limitation to the experimental use of such DFT, or to theoretical applications (*e.g.* obtaining NESS generalization of fluctuation-dissipation relations). A central result of our work is that generalized DFTs can be established *without relying on dual probabilities* by considering *joint* probability distributions for different complementary observables, instead of single PDFs. The joint probability distributions arise naturally from the above mentioned separation of two contributions to the total heat exchanged in a NESS. From this novel joint DFT all the known DFT and IFT follow in a straightforward way.

Among the fluctuation theorems formulated for Markov dynamics, the total trajectory entropy production  $\mathcal{S}[x; \sigma] = \ln(\mathcal{P}[x; \sigma]/\mathcal{P}[x^{\text{R}}; \sigma^{\text{R}}])$  plays a fundamental role [5, 6, 14]. Here  $\mathcal{P}[x; \sigma]$  ( $\mathcal{P}[x^{\text{R}}; \sigma^{\text{R}}]$ ) denotes the probability density of trajectory  $x$  (time reversed trajectory  $x^{\text{R}}$ ) in the forward (backward) protocol. We include in  $\mathcal{P}$  the initial distribution of  $x$ . We omit here-

after the final time  $\tau$  in all trajectory functionals, and use calligraphic symbols to denote functionals, and normal symbols to denote their values. It is straightforward to show that the total trajectory entropy production is odd upon time-reversal:  $\mathcal{S}[x^R; \sigma^R] = -\mathcal{S}[x; \sigma]$ . We show that any decomposition of  $\mathcal{S}$  in  $M$  distinct contributions,  $\mathcal{S}[x; \sigma] = \sum_i^M \mathcal{A}_i[x; \sigma]$ , each of them being odd  $\mathcal{A}_i[x^R; \sigma^R] = -\mathcal{A}_i[x; \sigma]$ , has a generating function satisfying the symmetry

$$\left\langle e^{-\sum_i^M \lambda_i \mathcal{A}_i[x; \sigma]} \right\rangle = \left\langle e^{-\sum_i^M (1-\lambda_i) \mathcal{A}_i[x; \sigma^R]} \right\rangle_R \quad (1)$$

where  $\lambda_i$  are arbitrary parameters and  $\langle \dots \rangle_R$  denotes average over trajectories in the reversed protocol. This symmetry is equivalent to the following generalized DFT for the joint probability of the  $\mathcal{A}_i[x; \sigma]$ 's:

$$\frac{P(A_1, A_2, \dots, A_M)}{P^R(-A_1, -A_2, \dots, -A_M)} = e^S \text{ with } S = \sum_{i=1}^M A_i \quad (2)$$

Note that the result involves no use of dual PDFs. To prove (1) we start by noting that the average of any observable  $\mathcal{O}[x]$  over trajectories satisfies the relation

$$\begin{aligned} \langle \mathcal{O}[x; \sigma] \rangle &= \int \mathcal{D}x \mathcal{P}[x^R; \sigma^R] \mathcal{O}[x; \sigma] e^{-\mathcal{S}[x^R; \sigma^R]} \\ &= \int \mathcal{D}x \mathcal{P}[x; \sigma^R] \mathcal{O}[x^R; \sigma] e^{-\mathcal{S}[x; \sigma^R]} \\ &= \langle \mathcal{O}[x^R; \sigma] e^{-\mathcal{S}[x; \sigma^R]} \rangle_R \end{aligned} \quad (3)$$

where we have used  $\mathcal{S}[x; \sigma] = -\mathcal{S}[x^R; \sigma^R]$  together with the change of variable  $x \rightarrow x^R$ . Considering  $\mathcal{S}[x; \sigma] = \sum_i^M \mathcal{A}_i[x; \sigma]$  the proof comes around  $\langle e^{-\sum_i^M \lambda_i \mathcal{A}_i[x; \sigma]} \rangle = \langle e^{-\sum_i^M \lambda_i \mathcal{A}_i[x^R; \sigma^R] - \mathcal{S}[x; \sigma^R]} \rangle_R = \langle e^{-\sum_i^M (1-\lambda_i) \mathcal{A}_i[x; \sigma^R]} \rangle_R$ . Eq.(2) is proved in a similar way, or also follows from (1) since it is a symmetry for the generating function of the joint distribution  $P(A_1, \dots, A_M)$ .

Before considering their particular application for explicit decompositions of  $\mathcal{S}$  we note that the symmetries (1) and (2) are valid for all times  $\tau$  for systems prepared in any initial condition. In particular, we see that the *total trajectory entropy production FTs*  $\langle e^{-\mathcal{S}} \rangle = 1$  and  $P(S)/P^R(-S) = e^S$  hold without further assumption.

We will consider two generic frameworks where our result applies: systems described (i) by Langevin dynamics, and (ii) by a Markov dynamics on discrete configurations, exemplifying their parallel features. Let us first consider a particle driven out of equilibrium by a constant force  $f$  in a potential  $U$ , subjected to the Langevin dynamics

$$\dot{x} = -\partial_x U(x; \alpha(t)) + f + \xi \quad (4)$$

where  $\alpha(t)$  represents a set of control parameters, and  $\xi(t)$  the Langevin noise,  $\langle \xi(t) \rangle = 0$ ,  $\langle \xi(t) \xi(t') \rangle = 2T \delta(t - t')$ , modelling the interaction of the system with a thermal bath at temperature  $T$ . We consider for simplicity a single degree of freedom  $x$ , but our results are easily generalized *e.g.* in larger dimensions and/or with more particles. For a stochastic trajectory generated by (4) we define the total exchanged heat as [17]

$$\beta \mathcal{Q}^{\text{tot}}[x; \sigma] = -\beta \int_0^\tau dt \dot{x} [\partial_x U(x; \alpha) - f] \quad (5)$$

The total exchanged heat in a trajectory can be split as  $\mathcal{Q}^{\text{tot}} = \mathcal{Q}^{\text{hk}} + \mathcal{Q}^{\text{ex}}$  [11]. Defining  $\phi(x; \sigma) = -\ln \rho_{\text{ss}}(x, \sigma)$  from the steady-state probability density  $\rho_{\text{ss}}(x, \sigma)$  at fixed values of  $\sigma = (\alpha, f)$  Hatano and Sasa [12] proposed

$$\beta \mathcal{Q}^{\text{hk}}[x; \sigma] = \int_0^\tau dt \dot{x} [\partial_x \phi(x; \sigma) - \beta (\partial_x U(x; \alpha) - f)] \quad (6)$$

for the housekeeping heat and

$$\beta \mathcal{Q}^{\text{ex}}[x; \sigma] = - \int_0^\tau dt \dot{x} \partial_x \phi(x; \sigma) \quad (7)$$

for the the excess heat. The Hatano-Sasa functional [12]  $\mathcal{Y}[x; \sigma]$  is then defined as

$$\mathcal{Y}[x; \sigma] \equiv \int_0^\tau dt \dot{\sigma} \partial_\sigma \phi(x; \sigma) = \beta \mathcal{Q}^{\text{ex}}[x; \sigma] + \Delta \phi(x; \sigma) \quad (8)$$

where  $\Delta \phi(x; \sigma) = \phi(x(\tau); \sigma(\tau)) - \phi(x(0); \sigma(0))$  is a time boundary term.

We now assume that the system is initially prepared in a NESS compatible with  $\sigma(0)$ . With the previous definitions of Eqs.(5),(6),(7) and (8) it is known that the total trajectory entropy production  $\mathcal{S}$  can be decomposed as the sum of two contributions, in two different ways [13]

$$\mathcal{S} = \mathcal{Y} + \beta \mathcal{Q}^{\text{hk}} = \Delta \phi + \beta \mathcal{Q}^{\text{tot}} \quad (9)$$

Let us now show that similar decompositions also exist for Markov dynamics: we consider discrete configurations  $\{\mathcal{C}\}$  with transition rates  $W(\mathcal{C} \rightarrow \mathcal{C}'; \sigma)$  between configurations. They depend on  $\sigma$ , an external control parameter which may vary in time. The probability density at time  $t$  obeys the Markov dynamics  $\partial_t P(\mathcal{C}, t) = \sum_{\mathcal{C}'} W(\mathcal{C}' \rightarrow \mathcal{C}; \sigma(t)) P(\mathcal{C}', t) - r(\mathcal{C}; \sigma(t)) P(\mathcal{C}, t)$  where  $r(\mathcal{C}; \sigma) = \sum_{\mathcal{C}'} W(\mathcal{C} \rightarrow \mathcal{C}'; \sigma)$  is the escape rate from configuration  $\mathcal{C}$ . An history of the system is described by the succession of configuration  $(\mathcal{C}_0, \dots, \mathcal{C}_K)$  visited by the system,  $\mathcal{C}_k$  being visited between  $t_k$  and  $t_{k+1}$ . Starting from initial distribution  $P_i(\mathcal{C}, \sigma)$ , the probability of an history is  $\mathcal{P}[\mathcal{C}; \sigma] = e^{-\int_0^\tau dt r(\mathcal{C}(t); \sigma(t))} \prod_{k=1}^K W(\mathcal{C}_{k-1} \rightarrow \mathcal{C}_k, \sigma_{t_k}) P_i(\mathcal{C}(0), \sigma(0))$  meaning that the mean value of an history-dependent observable  $\mathcal{O}$  is given by  $\langle \mathcal{O} \rangle = \sum_{K \geq 0} \sum_{\mathcal{C}_0 \dots \mathcal{C}_K} \int_0^t dt_K \dots \int_0^{t_2} dt_1 \mathcal{O}[\mathcal{C}, \sigma] \mathcal{P}[\mathcal{C}, \sigma]$ . We obtain that the total trajectory entropy production  $\mathcal{S}[\mathcal{C}; \sigma] = \ln(\mathcal{P}[\mathcal{C}; \sigma] / \mathcal{P}[\mathcal{C}^R; \sigma^R])$  has a first decomposition

$\mathcal{S} = \Delta\phi + \beta\mathcal{Q}^{\text{tot}}$  with  $\Delta\phi = \log \frac{P_1(\mathcal{C}(0), \sigma(0))}{P_1(\mathcal{C}(\tau), \sigma(\tau))}$  and

$$\beta\mathcal{Q}^{\text{tot}} = \sum_{k=1}^K \log \frac{W(\mathcal{C}_{k-1} \rightarrow \mathcal{C}_k, \sigma_{t_k})}{W(\mathcal{C}_k \rightarrow \mathcal{C}_{k-1}, \sigma_{t_k})} \quad (10)$$

Although there is no natural definition of  $\beta$  we keep the notation  $\beta\mathcal{Q}^{\text{tot}}$  to exemplify the parallel with Langevin dynamics.

Turning to the second decomposition, let's now assume that the initial distribution is steady-state:  $P_1 = \rho_{\text{ss}} = e^{-\phi}$ . One directly checks that the Hatano-Sasa functional  $\mathcal{Y}[\mathcal{C}, \sigma] = \int_0^\tau dt \dot{\sigma} \partial_\sigma \phi$  writes

$$\mathcal{Y} = [\phi(\mathcal{C}, \sigma)]_0^\tau - \sum_{k=1}^K [\phi(\mathcal{C}_k, \sigma_{t_k}) - \phi(\mathcal{C}_{k-1}, \sigma_{t_k})] \quad (11)$$

Besides, defining the house-keeping work as

$$\begin{aligned} \beta\mathcal{Q}^{\text{hk}}[\mathcal{C}, \sigma] &= \sum_{k=1}^K \log \frac{W(\mathcal{C}_{k-1} \rightarrow \mathcal{C}_k, \sigma_{t_k})}{W(\mathcal{C}_k \rightarrow \mathcal{C}_{k-1}, \sigma_{t_k})} \\ &+ \sum_{k=1}^K \phi(\mathcal{C}_k, \sigma_{t_k}) - \phi(\mathcal{C}_{k-1}, \sigma_{t_k}) \end{aligned} \quad (12)$$

we check that the decomposition  $\mathcal{S} = \mathcal{Y} + \beta\mathcal{Q}^{\text{hk}}$  holds [18]. The parallel between Markov and Langevin frameworks also appears by specializing to rates  $W(k \rightarrow k \pm 1) = e^{-\frac{\beta}{2}(V_{k \pm 1} - V_k)}$  of jumping on a 1d lattice from site  $k$  to  $k \pm 1$  in a tilted potential  $V_k = U_k - kf$ : in the continuum limit, one recovers the Langevin observables [18].

In the first decomposition,  $\mathcal{Y}$  can be identified with the so-called non-adiabatic contribution (since it vanishes for quasistatic protocols) to the trajectory entropy  $\mathcal{S}_{\text{na}} \equiv \mathcal{Y}$  while  $\beta\mathcal{Q}^{\text{hk}}$  (which is continuously produced in the steady-state) can be identified with the so-called adiabatic part  $\mathcal{S}_a \equiv \beta\mathcal{Q}^{\text{hk}}$  [16]. On the other hand, in the second decomposition of  $\mathcal{S}$ ,  $\Delta\phi$  can be identified with the system contribution  $\mathcal{S}_s \equiv \Delta\phi$  while  $\beta\mathcal{Q}^{\text{tot}}$  can be identified with the reservoir contribution  $\mathcal{S}_r \equiv \beta\mathcal{Q}^{\text{tot}}$  to the total trajectory entropy production.

The entropy decompositions of (9) satisfy the conditions for the application of the identities (1) or (2) since each term is odd with respect to time reversal in both decompositions. We can thus write DFTs (valid for all times  $\tau$ ) for the joint probabilities as

$$\frac{P(Y, \beta\mathcal{Q}^{\text{hk}})}{P^R(-Y, -\beta\mathcal{Q}^{\text{hk}})} = e^{Y + \beta\mathcal{Q}^{\text{hk}}} \quad (13)$$

$$\frac{P(\Delta\phi, \beta\mathcal{Q}^{\text{tot}})}{P^R(-\Delta\phi, -\beta\mathcal{Q}^{\text{tot}})} = e^{\Delta\phi + \beta\mathcal{Q}^{\text{tot}}} \quad (14)$$

with (14) valid for any initial distribution with  $\Delta\phi = \log \frac{P_1(\mathcal{C}(0), \sigma(0))}{P_1(\mathcal{C}(\tau), \sigma(\tau))}$  while (13) requires starting from the NESS. The corresponding IFTs write

$$\langle e^{-\lambda\mathcal{Y} - \kappa\beta\mathcal{Q}^{\text{hk}}} \rangle = \langle e^{-(1-\lambda)\mathcal{Y}_R - (1-\kappa)\beta\mathcal{Q}_R^{\text{hk}}} \rangle_R \quad (15)$$

$$\langle e^{-\lambda\Delta\phi - \kappa\beta\mathcal{Q}^{\text{tot}}} \rangle = \langle e^{-(1-\lambda)\Delta\phi_R - (1-\kappa)\beta\mathcal{Q}_R^{\text{tot}}} \rangle_R \quad (16)$$

Here  $\mathcal{X}_R$  denotes here  $\mathcal{X}[x; \sigma^R]$ ,  $\mathcal{X}$  representing  $\mathcal{Y}$ ,  $\mathcal{Q}^{\text{hk}}$ ,  $\mathcal{Q}^{\text{tot}}$  or  $\Delta\phi$ . From (13) and (14) we have, in terms of  $S_s$ ,  $S_r$ ,  $S_a$  and  $S_{\text{na}}$ , that  $P(S_s, S_r)/P^R(-S_s, -S_r) = e^{S_s + S_r}$  and  $P(S_a, S_{\text{na}})/P^R(-S_a, -S_{\text{na}}) = e^{S_a + S_{\text{na}}}$ . It is worth noting that these relations do not involve dual PDFs, and thus they can be tested for a physical system with a given dynamics. We also note that while one can show that  $S_a$  and  $S_{\text{na}}$  satisfy each one separately a DFT by using dual PDFs [16],  $S_s$  and  $S_r$  satisfy a joint DFT although they do not satisfy separately a DFT.

Let us now derive from an unified view the known FTs. As an intermediate step, it is useful to define a “dual” trajectory weight  $\mathcal{P}^\dagger[x]$  as [15, 16]

$$\mathcal{P}^\dagger[x; \sigma] = \mathcal{P}[x; \sigma] e^{-\beta\mathcal{Q}^{\text{hk}}[x; \sigma]}. \quad (17)$$

For Markov dynamics the dual probability  $\mathcal{P}^\dagger$  corresponds to the dynamics in the so-called dual rates  $W^\dagger(\mathcal{C} \rightarrow \mathcal{C}', \sigma) \equiv e^{-[\phi(\mathcal{C}', \sigma) - \phi(\mathcal{C}, \sigma)]} W(\mathcal{C}' \rightarrow \mathcal{C}, \sigma)$  which share the same steady state as the original dynamics. In the case of the Langevin dynamics of Eq. (4), it corresponds to trajectories generated by the equation  $\dot{x} = -\partial_x U^\dagger(x; \alpha(t)) + f^\dagger + \xi$  with  $U^\dagger = 2\phi/\beta - U$ ,  $f^\dagger = -f$ . This equation also has the same steady state as the original one. In both cases, let us stress that the dual dynamics corresponds to trajectories in a different physical system. We now follow (17) to write the dual joint PDF related to (13) as

$$P^\dagger(Y, \beta\mathcal{Q}^{\text{hk}}) = P(Y, -\beta\mathcal{Q}^{\text{hk}}) e^{\beta\mathcal{Q}^{\text{hk}}} \quad (18)$$

which is normalized. Integrating this relation over  $Y$ , we first obtain the DFT [16]  $P(\beta\mathcal{Q}^{\text{hk}}) = e^{\beta\mathcal{Q}^{\text{hk}}} P^\dagger(-\beta\mathcal{Q}^{\text{hk}})$ , and hence the IFT  $\langle e^{-\beta\mathcal{Q}^{\text{hk}}} \rangle = 1$  [13]. Using now successively (13) and (18)

$$\begin{aligned} P(Y) &= e^Y \int d(\beta\mathcal{Q}^{\text{hk}}) e^{\beta\mathcal{Q}^{\text{hk}}} P^R(-Y, -\beta\mathcal{Q}^{\text{hk}}) \\ &= e^Y \int d(\beta\mathcal{Q}^{\text{hk}}) P^{\dagger R}(-Y, \beta\mathcal{Q}^{\text{hk}}) \end{aligned} \quad (19)$$

we see that the DFT  $P(Y) = e^Y P^{\dagger R}(-Y)$  [15] holds. This implies the corresponding IFT  $\langle e^{-\mathcal{Y}} \rangle = 1$  [12] (also derived from setting  $\lambda = 1$ ,  $\kappa = 0$  in (15) and using the Speck-Seifert IFT).

As an illustration of our approach, let us show how joint FTs provide insights on the experimental error in the evaluation of entropy productions. Consider an experiment where the steady state can be evaluated for different values of the control parameter  $\sigma$ , *e.g.* microspheres optically driven in a liquid [19]. Having in hand an experimental evaluation  $\phi_{\text{exp}}$  of  $\phi$ , we write

$$\mathcal{S} = \mathcal{Y}_{\text{exp}} + \delta\mathcal{Y} + \beta\mathcal{Q}^{\text{hk}} \quad (20)$$

where  $\mathcal{Y}_{\text{exp}}[x; \sigma] = \int_0^\tau dt \dot{\sigma} \partial_\sigma \phi_{\text{exp}}$  and  $\delta\mathcal{Y} = \mathcal{Y} - \mathcal{Y}_{\text{exp}}$  is the difference between exact and experimental Hatano-Sasa functionals. Starting from the NESS associated to

$\sigma(0)$ , each of the terms in (20) is odd upon time-reversal, and we can use Eq. (2) for  $M = 3$ , which yields the DFT

$$P(Y_{\text{exp}}, \delta Y, \beta Q^{\text{hk}}) = P^{\text{R}}(-Y_{\text{exp}}, -\delta Y, -\beta Q^{\text{hk}}) e^{Y_{\text{exp}} + \delta Y + \beta Q^{\text{hk}}}.$$

Using (18), this gives  $P(Y_{\text{exp}}, \delta Y) = P^{\text{R}}(-Y_{\text{exp}}, -\delta Y) e^{Y_{\text{exp}} + \delta Y}$  and hence also the IFT

$$\langle e^{-\mathcal{Y}_{\text{exp}}} \rangle = \langle e^{-\delta \mathcal{Y}_{\text{R}}} \rangle_{\text{R}}^{\dagger}, \quad (21)$$

which allows to estimate the difference between the IFT  $\langle e^{-\mathcal{Y}} \rangle = 1$  and the experimentally obtained  $\langle e^{-\mathcal{Y}_{\text{exp}}} \rangle$ .

As a second example let us consider the situation in which the system, initially prepared in a NESS, has a variation of its parameters  $\sigma_i(t) = \sigma_i^0 + \delta\sigma_i(t)$  in such a way that  $|\frac{\delta\sigma_i(t)}{\sigma_i^0}| \ll 1$ , with  $\sigma_i^0 = \sigma_i(0)$  and  $\delta\sigma_i(0) = 0$ . In this context a modified Fluctuation-Dissipation Theorem has been recently derived in Ref. [20] which relates dissipation under small perturbations around a NESS, with fluctuations in the corresponding steady state. One can expand the exponentials in (15) up to second order in  $\delta\sigma$  and then in powers of  $\lambda$  and  $\kappa$ . To second order in  $\lambda$  and order zero in  $\kappa$ , we arrive to (see [18] for details),

$$\langle \mathcal{B}_{ij}(0, \tau) \rangle_0 = \langle \mathcal{B}_{ij}(\tau, 0) e^{-\beta \int_0^\tau dt \dot{x}(t) v_s(x(t); \sigma^0)} \rangle_0, \quad (22)$$

where  $\mathcal{B}_{ij}(t, t') = \frac{\partial \phi(x(t); \sigma^0)}{\partial \sigma_i} \frac{\partial \phi(x(t'); \sigma^0)}{\partial \sigma_j}$  and  $v_s = \beta^{-1} \partial_x \phi - (\partial_x U - f)$  corresponds to the average velocity in the NESS associated to  $\sigma$ . For systems with Boltzmann-Gibbs steady state, the obtained result reduces to the symmetry  $\mathcal{C}_{ij}(\tau) = \mathcal{C}_{ij}(-\tau)$ , with  $\mathcal{C}_{ij}(\tau) = \langle \mathcal{B}_{ij}(0, \tau) \rangle_0$ . Eq.(22) can also be derived from Eq.(19) in Ref.[15]. However, the use of the joint PDF can lead us to obtain further new results. Let us introduce a weighted correlation function as

$$\mathcal{C}_{ij}^W(\tau, 0) = \frac{\langle \mathcal{B}_{ij}(0, \tau) e^{-\frac{\beta}{2} \int_0^\tau dt \dot{x}(t) v_s(x(t); \sigma^0)} \rangle_0}{\langle e^{-\frac{\beta}{2} \int_0^\tau dt \dot{x}(t) v_s(x(t); \sigma^0)} \rangle_0} \quad (23)$$

This correlation function carries explicit information about the lack of detailed balance and reduces to the usual one when the system is able to equilibrate. Using Eq.(15) with  $\kappa = \frac{1}{2}$  and repeating the same procedure we have done in order to obtain Eq.(22) we arrive to the result  $\mathcal{C}_{ij}^W(\tau, 0) = \mathcal{C}_{ij}^W(0, \tau)$ , which is completely symmetric and reduces to the known result for equilibrium dynamics when detailed balance holds.

In conclusion, the identities (1) and (2) and their immediate consequences for Markovian systems are the main message of this work. Equations (1) or (2) indeed contain, as particular cases, several known FTs such as the ones previously derived by Hatano and Sasa [12], Speck and Seifert [13], Chernyak *et al.* [15] and Esposito and Van den Broek [16]. In addition, an exact DFT, valid for all times  $\tau$ , holds for the joint distribution of the reservoir and system entropy contributions to the total trajectory entropy production, although each contribution does

not do it separately, as given by Eq.(14). Also a similar DFT holds for the joint distribution of the adiabatic and nonadiabatic entropy contributions to the total trajectory entropy, as given by Eq.(13). It is worth to mention that for the type of NESS discussed here,  $M = 2$  decompositions of the total trajectory entropy production are obtained, Eq.(9), and thus two-variable joint PDFs are all that is needed for the corresponding DFTs. We have shown and example with  $M = 3$  for handling experimental errors in the Hatano-Sasa FT. In any case, in the light of (1) and (2), obtaining an adequate minimal  $M$ -decomposition of the total trajectory entropy production constitutes the cornerstone towards the derivation of generalized FTs for non-equilibrium systems.

This work was supported by CNEA, CONICET (PIP11220090100051), ANPCYT (PICT2007886), V.L. was supported in part by the Swiss NSF under MaNEP and Division II and thanks CNEA for hospitality.

- 
- [1] C. Bustamante, J. Liphardt, F. Ritort, *Physics Today*, **58** 43 (2005).
  - [2] F. Ritort, *Advances in Chemical Physics* **137**, 31 (2008) Ed. Wiley & Sons. arXiv:0705.0455v1.
  - [3] Denis J. Evans, E. G. D. Cohen and G. P. Morriss, *Phys. Rev. Lett.* **71**, 2401 (1993).
  - [4] G. Gallavotti and E. G. D. Cohen, *Phys. Rev. Lett.* **74**, 2694 (1995).
  - [5] J. Kurchan, *J. Phys. A: Math. Gen.* **31** 3719 (1998).
  - [6] J. L. Lebowitz and H. Spohn, *J. Stat. Phys.* **95** 333 (1999).
  - [7] C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997); C. Jarzynski, *Phys. Rev. E* **56**, 5018 (1997).
  - [8] G. E. Crooks, *J. Stat. Phys.* **90**, 1481 (1998); G. E. Crooks, *Phys. Rev. E* **61**, 2361 (2000).
  - [9] J. Kurchan, *J. Stat. Mech.* (2007) P07005
  - [10] R. J. Harris and G. M. Schütz, *J. Stat. Mech.* P07020
  - [11] Y. Oono and M. Paniconi, *Prog. Theor. Phys. Suppl.* **130**, 29 (1998)
  - [12] Takahiro Hatano and Shin-ichi Sasa, *Phys. Rev. Lett.* **86**, 3463 (2001).
  - [13] T. Speck and U. Seifert, *J. Phys. A: Math. Gen.* **38**, 581 (2005)
  - [14] U. Seifert, *Phys. Rev. Lett.* **95** 040602 (2005)
  - [15] V. Y Chernyak, M. Chertkov and C. Jarzynski, *J. Stat. Mech.* P08001 (2006).
  - [16] M. Esposito and C. Van den Broeck, *Phys. Rev. Lett.* **104**, 090601 (2010)
  - [17] K. Sekimoto, *Prog. Theor. Phys. Suppl.* **130**, 17 (1998)
  - [18] R. García-García, V. Lecomte, A. B. Kolton and D. Domínguez, in preparation.
  - [19] E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante and J. Liphardt, *PNAS* **101** 15038 (2004).
  - [20] J. Prost, J.-F. Joanny and J. M. R. Parrondo, *Phys. Rev. Lett.* **103** 090601 (2009)